

The Public-Signal Alpha Half-Life Hypothesis

Akhilesh Varma

Independent Researcher

akhverm@gmail.com

September 1, 2026

Abstract

This paper develops a theory of alpha decay for quantitative-finance signals whose construction is public, reproducible, and economically implementable. The central claim is deliberately narrower than the statement that every open-source model has zero alpha. A public signal may retain conditional or temporary predictive content, but its net implementable alpha should tend to decline as information diffuses, capital crowds the trade, arbitrage compresses mispricing, execution costs consume gross returns, and the market regime changes. We formalize this intuition with a public-signal state variable, a capital-crowding process, a cost function, and a regime-survival process. The resulting model yields an alpha half-life and a break-even time at which expected net alpha reaches zero. We derive propositions showing when public disclosure necessarily lowers net alpha and when alpha can survive through capacity, latency, non-public data, changing implementation, or structural compensation for risk. We then propose an empirical design based on publication-date event studies, rolling out-of-sample tests, survival analysis, multiple-testing corrections, and cost-aware portfolio construction. The paper does not assert that all Kalman filters, Bayesian models, or open-source strategies have no alpha; it argues that public availability is an economically meaningful mechanism that makes persistent, cost-adjusted alpha a demanding empirical claim.

Keywords: alpha decay; public signals; quantitative finance; crowding; arbitrage; transaction costs; backtest overfitting; market efficiency.

1 Introduction

Quantitative models are unusually easy to disseminate. A paper can publish an equation, a repository can publish executable code, and a data vendor can make the required inputs inexpensive to obtain. This reproducibility is beneficial for science, but it changes the economics of the signal. Once a method becomes observable and tradable, other participants can copy it, hedge against it, compete for the same liquidity, and remove the price discrepancy on which the signal depends.

The intuition is often expressed too strongly: a public Kalman filter, Bayesian model, or open-source trading rule has “no alpha left.” That universal statement is not defensible without a model-specific and time-specific test. Some public signals can survive because they compensate for risk, require scarce data, have limited capacity, depend on execution speed, are costly to arbitrage, or are conditional on a regime that has not disappeared. The appropriate claim is instead about a tendency and a mechanism.

This paper formalizes the following hypothesis:

For a reproducible public signal, the expected net implementable alpha has a finite half-life that shortens with information exposure, capital crowding, arbitrage intensity, and regime instability, unless new sources of differentiation or compensation offset the decay.

The distinction between gross predictive content and net alpha is essential. Let s_t be a signal-generated position and r_{t+1} the next-period asset return. Gross signal value may be written as

$$\alpha_t^{\text{gross}} = \mathbb{E}[s_t r_{t+1} \mid \mathcal{I}_t] - \mathbb{E}[s_t r_{t+1}^{\text{bench}} \mid \mathcal{I}_t]. \quad (1)$$

Net alpha is

$$\alpha_t^{\text{net}} = \alpha_t^{\text{gross}} - \text{Cost}_t - \text{RiskCharge}_t, \quad (2)$$

where costs include commissions, bid-ask spread, market impact, financing, borrow, turnover, and slippage. A model may retain a positive rank correlation with returns while having non-positive net alpha.

The paper makes four contributions. First, it introduces a compact alpha-decay process that separates public exposure, crowding, arbitrage, costs, and regime survival. Second, it derives conditions under which public disclosure lowers expected net alpha. Third, it explains why open-source model families should not be treated as a homogeneous group with identical remaining alpha. Fourth, it proposes a falsifiable empirical protocol that can distinguish publication decay from data mining, risk compensation, and changing market structure.

2 Related Evidence and Interpretation

McLean and Pontiff study 97 variables reported to predict cross-sectional stock returns. Their abstract reports that portfolio returns are 26% lower out of sample and 58% lower after publication, implying a 32% publication-informed decline; they also report stronger post-publication declines for predictors with higher in-sample returns and increased correlations among published predictor portfolios [1]. This is direct evidence consistent with investors learning from academic publication, although it does not imply that every signal becomes worthless.

Bowles, Reed, Ringgenberg, and Thornock study the timing of returns around the first public release of anomaly trading signals [2]. Brogaard studies anomaly decay by decomposing information in prices and reports that different anomaly types can respond differently to publication and information changes [3]. These findings motivate a heterogeneous model: public exposure is not the only determinant, and the decay path need not be universal.

Bailey and López de Prado show why a reported backtest can exaggerate the probability that a strategy will work out of sample. Their deflated Sharpe ratio addresses selection bias from multiple testing, backtest overfitting, and non-normality [4]. Therefore, a decline in observed alpha after publication can reflect both economic competition and statistical selection. The empirical design must address both.

3 A Model of Public-Signal Alpha Decay

3.1 Signal state

Consider a strategy with a reproducible signal z_t . Let $A_t \geq 0$ measure public exposure: the degree to which the signal's construction, features, code, and performance are accessible to potential imitators. Let $K_t \geq 0$ measure capital crowding, $H_t \geq 0$ measure arbitrage intensity, and $R_t \in [0, 1]$ measure the survival of the market regime under which the signal was discovered.

The latent gross alpha before implementation costs is

$$g_t = g_0 R_t \exp(-\lambda_A A_t - \lambda_K K_t - \lambda_H H_t) + \nu_t, \quad (3)$$

where g_0 is initial gross alpha, $\lambda_A, \lambda_K, \lambda_H \geq 0$ are decay sensitivities, and ν_t is a zero-mean innovation. The exponential form is not asserted as a universal law; it is a tractable reduced form in which each pressure scales the remaining edge.

Net alpha is

$$n_t = g_t - c_0 - c_1 \text{Turnover}_t - c_2 \text{Impact}(K_t) - c_3 \text{Borrow}_t, \quad (4)$$

with non-negative cost coefficients. A signal is economically viable only if $\mathbb{E}[n_t] > 0$ after costs and risk adjustment.

3.2 Public exposure and capital dynamics

Public exposure may increase after a paper, repository release, conference presentation, social-media dissemination, or commercial packaging:

$$A_{t+1} = (1 - \delta_A) A_t + \omega_t, \quad (5)$$

where $\omega_t \geq 0$ is new exposure and δ_A is forgetting or replacement by newer information. Capital responds with delay:

$$K_{t+1} = (1 - \delta_K) K_t + \chi A_t + \zeta_t, \quad (6)$$

where χ measures how public exposure attracts capital and ζ_t contains other capital-flow shocks. The delay allows a publication to precede the full crowding effect.

Arbitrage intensity can be modeled as

$$H_{t+1} = (1 - \delta_H) H_t + \phi K_t + \psi A_t, \quad (7)$$

where ϕ captures capital-enabled arbitrage and ψ captures information-enabled competition even before large capital arrives.

Regime survival follows

$$R_{t+1} = R_t \exp(-\lambda_R q_t), \quad (8)$$

where $q_t \geq 0$ measures structural change, such as changes in market participants, microstructure, regulation, technology, or the data-generating process.

3.3 Alpha half-life

If A_t, K_t, H_t are approximated locally as linear trends and R_t is locally constant, the deterministic gross-alpha component is

$$\bar{g}_t = G \exp(-\Lambda t), \quad (9)$$

where $G > 0$ is the local initial edge and

$$\Lambda = \lambda_A \dot{A} + \lambda_K \dot{K} + \lambda_H \dot{H}. \quad (10)$$

When $\Lambda > 0$, the gross-alpha half-life is

$$t_{1/2} = \frac{\log 2}{\Lambda}. \quad (11)$$

The net-alpha break-even time solves

$$G \exp(-\Lambda t^*) = c, \quad (12)$$

so that

$$t^* = \frac{1}{\Lambda} \log \left(\frac{G}{c} \right), \quad G > c. \quad (13)$$

If $G \leq c$, the strategy is not net profitable even before decay.

4 Propositions

Proposition 1 (Public exposure lowers gross alpha). *Suppose $g_0 > 0$, $R_t > 0$, and $\lambda_A > 0$. Holding K_t , H_t , and R_t fixed, $\partial g_t / \partial A_t < 0$.*

Proof. Differentiating equation (3) gives

$$\frac{\partial g_t}{\partial A_t} = -\lambda_A g_0 R_t \exp(-\lambda_A A_t - \lambda_K K_t - \lambda_H H_t) < 0.$$

□

Proposition 2 (Crowding can eliminate net alpha). *Suppose $\lambda_K > 0$, $K_t \rightarrow \infty$, and the cost function in equation (4) is non-decreasing in K_t . Then $\limsup_{t \rightarrow \infty} \mathbb{E}[n_t] \leq 0$ under equation (3).*

Proof. The exponential term involving K_t converges to zero as $K_t \rightarrow \infty$. Thus expected gross alpha converges to zero or below after any non-positive risk adjustment. Since implementation cost is non-negative and non-decreasing in crowding, expected net alpha cannot have a positive limiting value. □

Proposition 3 (Public alpha can survive). *Positive expected net alpha can persist if at least one of the following holds: $\lambda_A = 0$ for the relevant information, K_t is capacity-limited, H_t is constrained by limits to arbitrage, R_t is maintained by a persistent risk compensation, or the strategy continually changes in a way not captured by the public signal state.*

Proof. Equation (4) permits positive net alpha whenever the gross component exceeds costs and risk charge. Each listed condition prevents the exponential gross-alpha component from converging to zero or prevents costs from rising enough to offset it. The proposition is sufficient, not necessary. □

5 Why Model Families Do Not All Lose Alpha Equally

A Kalman filter is a general estimation framework, not one unique trading signal. A Bayesian model is a broad inferential family, not a single portfolio rule. Open-source availability therefore does not identify A_t , K_t , H_t , or R_t by itself. The same public model can have different decay profiles depending on its inputs, latency, capacity, execution, and economic source of return.

The correct unit of analysis is therefore not “the Kalman filter” or “Bayesian models” in the abstract. It is a fully specified strategy: data, feature construction, training window, forecast horizon, portfolio rule, execution schedule, costs, leverage, capacity, and publication history.

Table 1: Mechanisms that can accelerate or slow public-signal alpha decay.

Mechanism	Expected effect
Public code and paper	Raises exposure and lowers replication cost
Low-latency implementation	Can preserve edge for fast operators but increases infrastructure competition
High capacity	Makes crowding and market impact more important
Low liquidity	Can leave mispricing but makes costs and limits to arbitrage severe
Risk compensation	Can survive publication if the return is payment for bearing systematic risk
Non-public or expensive data	Keeps the effective signal private even if the model family is public
Regime change	Can destroy the relationship that generated the historical signal
Signal mutation	Can create a new signal, but may also introduce overfitting

Table 2: Minimum tests for the public-signal alpha-decay hypothesis.

Test	Requirement
Publication event study	Define first public exposure and estimate pre/post performance with leads and lags
Rolling evaluation	Re-estimate performance through time without using future data
Cost-aware returns	Include spread, impact, turnover, borrow, financing, and delay
Capacity analysis	Report performance as capital and participation increase
Multiple testing control	Use deflated Sharpe, reality-check, or related corrections
Risk decomposition	Separate risk compensation from abnormal return
Heterogeneity analysis	Compare liquidity, asset class, horizon, signal type, and implementation speed
Replication test	Reimplement the published rule independently and archive all choices

6 Empirical Research Design

The central empirical test is whether a signal’s cost-adjusted performance changes around public exposure while controlling for statistical selection and market conditions. A practical panel design is

$$y_{i,t} = \alpha_i + \tau_t + \beta_1 \text{PostPublication}_{i,t} + \beta_2 \text{Crowding}_{i,t} + \beta_3 \text{Cost}_{i,t} + \Gamma X_{i,t} + \varepsilon_{i,t}, \quad (14)$$

where $y_{i,t}$ is out-of-sample alpha or a cost-adjusted performance measure for signal i at time t , α_i and τ_t are signal and time effects, and $X_{i,t}$ includes liquidity, volatility, market trend, and regime controls.

A credible protocol should use the following design choices.

The null hypothesis should not be “all public models have no alpha.” A sharper null is H_0 : after controlling for costs, risk, multiple testing, and regime changes, the post-publication decline is no larger than the decline for comparable non-public or not-yet-published signals. The alternative predicts a negative post-publication effect that increases with initial performance, replication ease, and crowding exposure.

7 Discussion

The theory explains why a model can remain statistically interesting while becoming economically unusable. A forecast probability may retain a small correlation with returns, but a small edge can be dominated by turnover, spread, impact, financing, and estimation error. Similarly, a model can work in a narrow regime while losing value after market participants adapt.

The framework also clarifies why open-source publication is not inherently destructive. Scientific openness can reduce duplication, improve implementation, and reveal errors. It may also move a strategy from a private information advantage toward a public risk premium. In that case, publication changes who can access the compensation but does not necessarily eliminate the compensation.

There are several alternative explanations for apparent alpha decay. A published strategy may have been overfit before publication. Data definitions may change. A portfolio may become less liquid. The market may experience a structural break unrelated to the paper. Researchers may selectively publish positive strategies. These explanations are why the paper treats publication as a mechanism to test, not as an automatic causal conclusion.

8 Limitations

The exponential decay law is a reduced-form model, not a universal law of markets. Exposure, crowding, arbitrage, and regime change are difficult to measure. Publication dates may be endogenous because researchers publish when a signal is strong. Capital-flow measures may be noisy. Some signals are difficult to replicate even when their broad idea is public, and some apparent alpha is a risk premium rather than mispricing.

The framework does not prove that any named model family has zero current alpha. Such a claim would require a defined strategy, a data vintage, an implementation rule, a cost model, and a time-stamped evaluation. It also cannot infer future alpha from historical decay alone.

9 Conclusion

Public quantitative models should not be described as automatically worthless, but neither should historical backtests be treated as permanently available alpha. The theory in this paper predicts a finite half-life for public-signal alpha when information exposure attracts capital, capital intensifies arbitrage, and implementation costs rise. The same theory permits persistent alpha when capacity is limited, arbitrage is constrained, data remain scarce, the return compensates for risk, or implementation continues to evolve.

The useful implication is methodological. The relevant question is not whether a Kalman filter, Bayesian model, or open-source algorithm has “alpha left” in the abstract. The relevant question is whether a fully specified, cost-adjusted, out-of-sample strategy retains economically meaningful performance after publication, crowding, multiple-testing correction, and regime controls. That question is testable, and the proposed model provides a structure for testing it without overstating what the evidence can show.

References

- [1] McLean, R. D. and Pontiff, J. *Does Academic Research Destroy Stock Return Predictability?* Journal of Finance 71, 2016. DOI: <https://doi.org/10.1111/jofi.12365>. Working-paper record: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2156623.
- [2] Bowles, B., Reed, A. V., Ringgenberg, M., and Thornock, J. *Anomaly Time*. SSRN working-paper record: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3069026.
- [3] Brogaard, J. *What Drives Anomaly Decay?* SSRN working-paper record: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4934042.
- [4] Bailey, D. H. and López de Prado, M. *The Deflated Sharpe Ratio: Correcting for Selection Bias, Backtest Overfitting, and Non-Normality*. Journal of Portfolio Management 40(5), 2014, pp. 94–107. DOI: <https://doi.org/10.3905/jpm.2014.40.5.094>.